



## A Corpus Based Description of Urdu Affixes, A Morphological Perspective

Barira Tanveer \*

Abrar Hussain Qureshi\*\*

Mehmood ul Hassan\*\*\*

Corpus based morphology has emerged into a science across the languages and in various specialized domains of applied linguistics. However, the job of a morphologist is not easy, as there are many complex languages that have significant effect of foreign elements. (Hartman, 1983). His task becomes all the more difficult when he handles one of the most inflectional and derivational languages as Urdu that flourished under the impact of many regional languages. The undertaken research is exploratory in nature. The focus has been to determine the corpus based description of Urdu affixes from morphological perspective. There is a long range of Urdu affixes, originating from different languages with various functions (Bogel, Butt and Sulger 2008). The proposed research is expected to help Urdu lexicologists, morphologists, lexicographers, etc. to handle the nature of Urdu affixes in a potential way for various linguistic uses.

**Key Words:** Description, affixes, bilingual, Urdu-English, lexicography,

### Introduction

One of the major ways to handle a language is mastering the process of its morphology. Bauer (1998) asserts that as lexis and grammar are present on one continuum of language, mastering word formation can improve overall proficiency of the language. Morphology of Urdu language can be described on the same patterns as English because both languages developed under the impact of many regional languages. The patterns of descriptions are suppletion, clipping, conversion, acronym, blends, compounding and so on. Affixation is one of the major techniques of forming new words in Urdu language.

<sup>1,2</sup>University of Sahiwal.

<sup>3</sup>KFUET

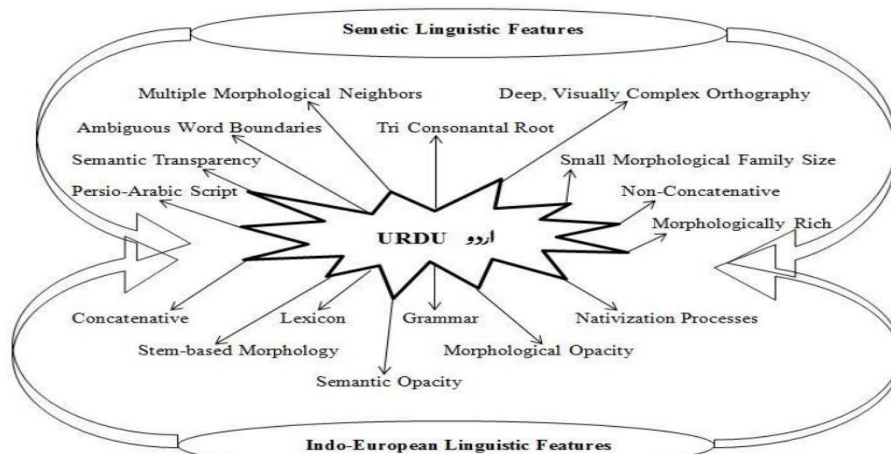


Fig. 1 Architecture depicting linguistic features of Urdu language



According to Butt and King (2005) most Urdu words are made of a root that is commonly a semantic carrier. Affixes are added to these roots to create a wide range of new words with syntactic and semantic information. This process is called affixation. The affix that is added in the beginning of an Urdu word is called “prefix” while the affix that is added to the end of an Urdu root is called “suffixes”. The base word occupies the central place in this process of affixation.

Determining lexical relations have been one of the dominating interests of Urdu-English bilingual lexicographers (McCarthy, 2002). It is assumed that total lexicon of a language is linked with a single semantic network. It is not an easy job for a bilingual lexicographer to define a word, keeping its identification with the words with which it is semantically related. Affixes have meanings in Urdu. They are not simply the combinations of two or three letters. That is why, they are described as morphemes. A “morpheme” is defined as the smallest unit of grammar. Sometimes, these morphemes change the meanings.

### Prefixes

Prefixes operate on many dimensions in the Urdu language. They belong to different groups with specific functions. According to Sabzwari (1956), the origin of most of the prefixes is not indigenous. They have been borrowed from various languages of the region as Urdu had been under the impact of many languages. Some prefixes have been modified for the indigenous use while some borrowed prefixes have been used without any modification. They sustain their original condition. Some prefixes of Urdu have been imported from Sanskrit. A few of Sanskrit prefixes have been changed with the passage of time. Most of the in use prefixes of the Urdu language are following: (Sabzwari, 1956)

#### 1. (نا) *na*

پسند (liking)

لائق (intelligent)

خدا (God)

کام (job)

نا پسند (unliking)

نا لائق (unintelligent)

نا خدا (boat man)

نا کام (unsuccessful)



Word	Cooccurrences	Candidates	T-score	MI	↓ LogDice
1 انصافی	225	361	14.99	11.46	9.06 ...
2 ممکن	377	12,006	19.28	7.14	8.92 ...
3 چا	194	1,460	13.91	9.23	8.73 ...
4 ابلی	175	194	13.23	11.99	8.71 ...
5 آشنا	143	1,043	11.94	9.27	8.33 ...
6 پسند	262	15,342	15.98	6.27	8. Back to the or

The prefix (نا) *na* has a negative meaning. It is very productive in the Urdu language (Haq, 1958). . It is attached with nouns and adjectives. After the attachment of prefix, some bases function as adjectives and some as nouns. The third example is functioning as noun while the rest of the examples are functioning as adjectives. The prefix (نا) creates little morphophonemic alterations in the base.

2. .(بے) *bay*

ایمان	بے ایمان	(faithless)
ڈھب	بے ڈھب	(formless)
غیرت	بے غیرت	(honorless)
داغ	بے داغ	(spotless)



Word	Cooccurrences	Candidates	T-score	MI	↓ LogDice
1 نظير	3,397	4,058	58.22	9.74	10.77 ...
2 شك	2,670	7,812	51.52	8.44	10.34 ...
3 شمار	2,325	7,457	48.07	8.31	10.15 ...
4 حد	2,140	13,680	45.97	7.32	9.90 ...
5 گناہ	1,501	8,445	38.53	7.50	9.50 ...
6 پناہ	1,198	5,530	34.46	7.79	9. ...

The prefix (بے) *bay* has a negative meaning. It is very productive in the Urdu language. It is attached with nouns. After the attachment of prefix, the base functions as adjective. It creates little morphophonemic alterations in the base.

### 3. (ن) *noon*

کام	نکما	(without work)
ڈر	نڈر	(fearless)
	نرالا <i>nerala</i>	(novel)

The prefix (ن) *noon* is very productive in the Urdu language and is attached with a number of bases. It is a negative prefix and creates negative meanings. It triggers morphophonemic alternations in the base. It is attached with nouns and after the prefix, the base functions as adjective.

### 4. (شاه) *shah*

باز	شہباز	(falcon)
راہ	شاہراہ	(highway)
سوار	شہسوار	(princely)
نواز	شاہ نواز	(great generous)



Word	Cooccurrences	Candidates	T-score	MI	↓ LogDice	
1 محمود	1,219	6,401	34.87	9.54	10.83	...
2 رخ	636	4,693	25.17	9.05	10.01	...
3 ولی	428	3,373	20.65	8.95	9.53	...
4 فیصل	293	4,051	17.06	8.14	8.93	...
5 صاحب	680	41,990	25.66	5.98	8.60	...
6 غازی	190	1,651	13.75	8.81	8.51	Back to t

The prefix (شاه) *shah* has the meaning of great. It is very productive in the Urdu language. Mostly, it is attached with nouns and occasionally with adjectives as in the last example. After the attachment of prefix, the base functions sometimes as adjective and sometimes as noun as in شاه راہ *shah-raah* (highway).

#### 5.(بن) *ben*

Word	Cooccurrences	Candidates	T-score	MI	↓ LogDice	
4 مسعود	336	2,454	18.22	7.40	7.73	...
5 عباس	345	4,751	18.37	6.49	7.71	...
6 عمر	428	18,967	19.95	4.80	7.68	...
7 علی	535	36,090	21.87	4.19	7.68	...
8 مالک	331	7,057	17.88	5.86	7.59	...
9 سکتے	407	22,530	19.27	4.48	7.54	...

The prefix (بن) is very productive in the Urdu language and is attached with a number of nouns. It is also a negative prefix and creates negative meanings. It does not trigger



morphophonemic alternations in the base. It is attached with nouns and after the prefix, the base functions as adjective.

6.( ہم ) ham

	آغوش	ہم آغوش	(embracing)				
	آواز	ہم آواز	(harmonious)				
	پیشہ	ہم پیشہ	(of the same trade)				
	نوا	ہم نوا	(group singing)				
6	آہنگی	828	865	28.72	9.06	7.94	...
7	جنس	639	3,734	25.01	6.57	7.53	...
8	ان	2,172	289,344	35.45	2.06	7.48	...
9	اپنی	1,105	95,738	28.07	2.68	7.46	...
10	دونوں	731	27,656	25.20	3.88	7.45	...
11	اس	3,951	658,961	44.02	1.74	7.40	...
12	آہنگ	511	785	22.54	8.50	7.25	Back t

The prefix *ham* (ہم) has the meaning of mutual. It is very productive in the Urdu language. Mostly, it is attached with nouns. After the attachment of prefix, the base functions as adjective. It creates little morphophonemic alterations in the base.

7.( ذی ) zi

	ہوش	ذی ہوش	(wise)
	شعور	ذی شعور	(intelligent)
	اختیار	ذی اختیار	(one in authority)

Word	Cooccurrences	Candidates	T-score	MI	↓ LogDice		
1	الحجہ	3	323	1.73	15.94	8.21	...
2	ا	3	39,049	1.73	9.02	1.33	...



The prefix *zi* (ذی) has the meaning of possession. It is very productive in the Urdu language. Mostly, it is attached with nouns. After the attachment of prefix, the base functions as adjective. It creates little morphophonemic alterations in the base.

8. (پاک) *paak*

Word	Cooccurrences	Candidates	T-score	MI	↓ LogDice	
1 فضائیہ	3	725	1.73	10.56	6.78	...
2 فوج	7	19,125	2.63	7.06	3.57	...
3 چین	3	9,616	1.72	6.83	3.33	...
4 بمیں	4	25,615	1.96	5.83	2.35	...
5 صلی	3	34,118	1.68	5.00	1.52	...
6 بو	13	229,018	3.43	4.37	0.89	Back to the

The prefix (پاک) has a positive meanings of pure. It is not very productive in the Urdu language. It is attached with a limited number of bases. It is attached with nouns and after the attachment of prefix, the base functions as adjective. It creates little morphophonemic alterations in the base.

### Suffixes

The case of suffixes is not much different. Most of the Urdu suffixes are borrowed over a long span of time. They also belong to different languages. Some of the most frequently used Urdu suffixes are following:

1. ناک (naak)  
خوف (dreadful)  
بول (dreadful)  
ہیبت (dreadful)  
شرمناک (shameful)



Word	Cooccurrences	Candidates	T-score	MI	↓ LogDice
1 خطر	11	387	3.32	14.94	9.67 ...
2 اذیت	3	1,225	1.73	11.40	6.26 ...

The suffix (ناک) has a negative meaning. It is not very productive in the Urdu language. It is attached with a limited number of bases. It is attached with nouns and after the attachment of suffix, the base functions as adjective. It triggers little morphophonemic alterations in the base.

## 2. ( آ ر ) aar

گفت	گفتار	(conversation)
گرفت	گرفتار	(arrested)
سونا	سنار	(gold smith)
چمرا	چمار	(tanner)

The suffix ( آ ر ) is productive in the Urdu language in the sense that it is attached with a number of bases. It is attached with nouns and verbs. After the attachment of suffix, the base functions always as noun. It triggers morphophonemic alterations in the base.

## 3.( دان ) daan

آتش	آتش دان	(fireplace)
نمک	نمک دان	(saltcellar)
پان	پان دان	(betel-leaf box)
ریاضی	ریاضی دان	(Mathematician)
روشن	روشن دان	(Ventilator)

The suffix ( دان ) is very productive in the Urdu language. It is attached with a number of bases. It is attached with nouns and after the attachment of suffix, the base functions as noun . It triggers little morphophonemic alterations in the base.

## 4.( ستان ) staan

گل	گلستان	(garden)
بہار	بہارستان	(a place with the bloom of spring)
ریگ	ریگستان	(desert)
نخل	نخلستان	(oasis)





اد 95	ن ستا	کی نسی کھای رہما	سعی رواف
ادّ ہے۔ دوسرا مگر قصہ کا اس،	ن ستا	پہلے ہم تو یہ کی نسی کھای رہما	د
بندو۔ تھے ر شکا کا افلاس نی حا	ن ستا	کی تاؤں دیو کروڑ تینتیس بنود کے	ا
وزیر لی شما، یر وزن مافر کٹر	ن ستا	لوگ 7 کنندہ تجویز	ا
و ہند اور یر سپین نے ن مسلما پھد	ن ستا	کن کی مت حکومت عریل طو یر	ا
و ہند تقسیم تھے۔ مسلم غیر والے	ن ستا	ید کوں نو مسلما نے کس وقت کے	ا
بندو نے ں جنہو، بے ٹی بو صل	ن ستا	ابما عرب و چین: لکھانہ تراکا	ا
بندو بے د جو مو اثر ں یا نما یر م	ن ستا	ک ج آکا عر شاہڑے ایک اور میں	ا

The suffix (ستان) is very productive in the Urdu language. It is a Persian suffix. It is attached with a number of bases. It is attached with nouns and after the attachment of suffix, the base functions as noun. The suffix generates the meaning of enlargement. It triggers little morphophonemic alterations in the base.

#### 5. (سار) saar

ملن	ملنسا ر	(sociable)
شرم	شرمسا ر	(embarrassed)
خاک	خاکسا ر	(humble)

The suffix (سار) has a meaning of possessing some quality. It is not very productive in the Urdu language. It is attached with a limited number of bases. It is attached with nouns and after the attachment of suffix, the base functions as adjective. It triggers little morphophonemic alterations in the base.

#### 6. (گر) ger

بازی	بازی گر	(a conjurer)
کاری	کاری گر	(skilled worker)
چارہ	چارہ گر	(helper)



Word	Cooccurrences	Candidates	T-score	MI	↓ LogDice
1 جلوہ	271	967	16.46	11.82	10.61 ...
2 چارہ	96	1,505	9.79	9.69	8.98 ...
3 جادو	82	1,594	9.04	9.38	8.74 ...
4 کوزہ	33	120	5.74	11.79	7.81 ...
5 طیارہ	36	949	5.99	8.94	7.70 ...
6 مکانات	35	1,146	5.90	8.62	7.61 ...

The suffix (گر) has a meaning of possessing some quality. It is not very productive in the Urdu language. It is attached with a limited number of bases. It is attached with nouns and after the attachment of suffix, the base again functions as noun. It triggers little morphophonemic alterations in the base.

### 7.(ا) alif

جھوٹ	جھوٹا	(liar)	
گنج	گنجا	(bald)	
سچ	سچا	(true)	
گند	گندا	(bad)	
پیا سا	پیا سا	(thirsty)	
کافی پانی یہاں ہیں۔ سکتے پڑھ	<b>گندا</b>	باعث کے ت تس ج ہا ت بار بو	
کافی یہ نسبت کی شہروں کے م	<b>گندا</b>	ی کی مصر اہل ہوا۔ محسوس سا	
ہی ویسا تھی۔ مشابہ سے شاہ سیا	<b>گندا</b>	بے بنے پر کناروں کے اس، نالہ	
کافی علقہ یہ نسبت کی مقامات ت	<b>گندا</b>	نے تین علقے اس تھا۔ رہا لگ سا	
خاطر تیری کو اس میں اور بے	<b>گندا</b>	ے متعلق کے اس میں ہوں رہا کر	
کو معاشرے تو پورے ہو جائے	<b>گندا</b>	و گندکی کی جس کہ ہے کر دیتی	
کر بڑھ سے سب نے چیزوں جز	<b>گندا</b>	اری یہ ایک سے میں بے، ان کیا	
بدلے کے (ف ۶) ستھرے اور ا	<b>گندا</b>	نے مال کے ان اور (ف ۷) لو نہ	
بھی جو کہ بے گیا لگ ہاتھ جواز	<b>گندا</b>	سلام کیونکہ کرو آئے میں دل کام	



The suffix (ا) is very productive in the Urdu language. It is attached with a number of bases. It is attached with nouns. After the attachment of suffix, the base functions as adjective. It triggers little morphophonemic alterations in the base.

8.(مند) *mand*

عقل	عقل مند	(wise)
ضرورت	ضرورت مند	(needy)
حاجت	حاجت مند	(needy)
خرد	خرد مند	(wise)

Word	Cooccurrences	Candidates	T-score	MI	↓ LogDice	
1 صحت	744	6,922	27.26	10.46	11.04	...
2 خوابش	401	6,055	20.00	9.77	10.26	...
3 فائدہ	487	8,416	22.04	9.57	10.26	...
4 غیرت	218	1,794	14.76	10.64	10.12	...
5 بہرہ	153	601	12.37	11.71	9.91	...
6 دولت	215	4,896	14.64	9.17	9.53	Back to the c

The suffix (ندی) is very productive in the Urdu language. It is attached with a number of bases. It is attached with nouns. After the attachment of suffix, the base functions as adjective. It triggers little morphophonemic alterations in the base.

9. (آنی) *yee*

صاف	صفا نی	(cleanliness)
چڑھنا	چڑھانی	(ascend)
لڑنا	لڑانی	(quarrell)
پڑھنا	پڑھانی	(study)

The suffix (آنی) is very productive in the Urdu language. It is attached with a number of bases. It is attached with verbs and after the attachment of suffix, the base functions as noun. It triggers morphophonemic alterations in the base. The consonant sound ن is deleted from the base.



10. (واں) *waan*

پانچ	پانچواں	(fifth)
سات	ساتواں	(seventh)
آٹھ	آٹھواں	(eighth)
نو	نواں	(ninth)

The suffix (واں) is very productive in the Urdu language. It is attached with numerals. After the attachment of suffix, the base functions as adjective. It triggers little morphophonemic alterations in the base.

11.(بان) *baan*

شتر	شتر بان	(A camel keeper)
گاڑی	گاڑی بان	(A cart keeper)
غلہ	غلہ بان	(grain keeper)
سائے	سائے بان	(

The suffix (بان) is not very productive in the Urdu language. It is attached with a number of bases. It is attached with nouns and after the attachment of suffix, the base functions as noun. The suffix generates the meaning of an agent. It triggers little morphophonemic alterations in the base. The affixes that are used in the process of derivation in the Urdu language move in and out of the use by the native speakers. Some derivational affixes are in vogue for a bit of time and then their productivity decrease over a period of time as کج *kaj* in کج فہم *kaj-fahem* (insane fellow) or ہفت *haft* in ہفت پیکر *haft-paeker* (multi-dimensional character). Some derivational affixes have very significant position and they continue to be the part of the living language. Some Urdu affixes that are frequently used in the Urdu derivational process, are so all pervasive that native speakers start attaching those prefixes even with those roots where it does not look appropriate to do so e.g. گھر ساختہ *gher saakh-taa* in ساختہ *saakh-taa* (made in house).

Sometimes, three or four parallel affixes are being operated by the native speakers for the same purpose but they are not easily interchanged. This is a very complex aspect of Urdu derivational process. In this regard, the most common example is the negative affixes that carry negative meanings with the specific roots but they can not be interchanged with various roots. The native speakers of the Urdu language say لا فانی *la-fanni* (immortal), ناسمجھ *na-samegh* (foolish), بے شک *bey-shek* (undoubtedly), بن صلہ *bin-seela* (unsewn), انتھک *un-thak* (untiring), and نڈر *ni-der* (fearless), etc.

**Concluding Remarks:**

The undertaken research clearly shows the complex nature of Urdu affixes that belong to different languages and with various semantic functions. It also establishes the key role of affixes in Urdu morphology. Some are only responsible for Urdu inflections, added to make plural of



nouns. Some Urdu affixes only change the tense while some affixes are added to make the comparative and superlative degrees of the Urdu adjectives. Hook asserts (1974) there are some affixes that are class changing, meaning to say that change the class of the part of speech while some are class maintaining and perform some semantic and syntactic functions with the root. There is a long list of Urdu affixes. Serious efforts have been done to describe maximum patterns of Urdu affixes but in the end all the dimensions of Urdu affixes may not have been explored. The undertaken research motivates the Urdu lexicologists to further explore the delicacies of Urdu affixes from morphological and lexicographic point of view.

### References

- Baeuer, L.(1998). Vocabulary. London and New York: Routledge.
- Böogel, T., M. Butt and S. Sulger. (2008). Urdu Ezafe and the Morphology-Syntax Interface. Proceedings of LFG08.
- Butt, M. and T.H. King. 2005. The Status of Case. In V.Dayal and A.Mahajan (eds.) Clause Structure in South Asian Languages. Kluwer Academic Publishers, 153-198
- Haq, M. A. (1958). Qwaaed Urdu. Lahore Academy Lahore
- Hartman, R.R.K. (1983). On Theory and Practice. Lexicography Principles and Practice, Academic Press London & New York
- Hook, P. E. 1974. Compound Verbs in Hindi. The University of Michigan. USA
- MacCarthy,A.N.(2002) An Introduction to English Morphology, Edinbergh University Press.
- Sabzwari, S (1956) Urdu Zubaan ka Irteaqa. Dhaka Ghewera-e-Adab